

日本国特許庁  
JAPAN PATENT OFFICE

23.06.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

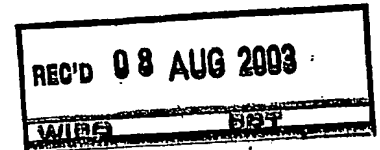
This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2002年11月 1日  
Date of Application:

出願番号 特願2002-319757  
Application Number:

[ST. 10/C]: [JP2002-319757]

出願人 松下電器産業株式会社  
Applicant(s):

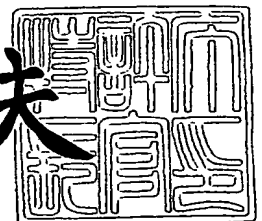


**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2003年 7月25日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



BEST AVAILABLE COPY

【書類名】 特許願

【整理番号】 2054041312

【提出日】 平成14年11月 1日

【あて先】 特許庁長官殿

【国際特許分類】 H04N 5/91

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 安方 満

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 小林 正明

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 松井 謙二

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 桑野 裕康

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 遠藤 充

【発明者】

【住所又は居所】 大阪府門真市大字門真 1006 番地 松下電器産業株式会社内

【氏名】 酒井 啓行

## 【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 下田代 雅文

## 【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

## 【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

## 【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

## 【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

## 【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 メタデータ制作装置及び制作方法

【特許請求の範囲】

【請求項1】 静止画コンテンツに関連するメタデータ的制作装置であって、  
音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、

上記静止画コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換するとともに、上記静止画コンテンツの各々の番地を上記メタデータに関連付けることを特徴とするメタデータ制作装置

。

【請求項2】 静止画コンテンツに関連するメタデータ的制作装置であって、  
音声入力手段と音声認識手段とメタデータ制作手段と上記静止画コンテンツに関連した辞書とを具備し、

上記静止画コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記静止画コンテンツに関連した辞書に関連付けて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換するとともに、上記静止画コンテンツの各々の番地を上記メタデータに関連付けることを特徴とするメタデータ制作装置。

【請求項3】 静止画コンテンツに関連するメタデータ的制作方法であって、  
音声入力手段と、音声認識手段と、メタデータ制作手段とを用いて、

上記静止画コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ制作手段によりメタデータに変換するとともに、上記静止画コンテンツの各々の番地を上記メタデータに関連付けることを特徴とするメタデータ制作方法。

【請求項4】 静止画コンテンツに関連するメタデータ的制作方法であって、  
音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを用いて、

上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付け

て認識し、上記メタデータ制作手段によりメタデータに変換するとともに、上記静止画コンテンツの各々の番地を上記メタデータに関連付けることを特徴とするメタデータ制作方法。

#### 【発明の詳細な説明】

##### 【0001】

#### 【発明の属する技術分野】

本発明は、静止画コンテンツ制作におけるメタデータ制作システム及び方法に関するものである。

##### 【0002】

#### 【従来技術】

近年、静止画コンテンツの制作において、これらコンテンツに関連したメタデータの付与することがおこなわれている。

##### 【0003】

しかしながら、上記メタデータの付与は、制作された静止画コンテンツのシナリオあるいはナレーション原稿をもとに、手作業でコンピュータ入力することにより制作する方法が一般的であり、相当な労力の必要な方法であった。

##### 【0004】

#### 【特許文献1】

特開平9-130736号公報

##### 【0005】

#### 【発明が解決しようとする課題】

本願発明は、上記従来の問題点に係る課題を解決することを目的とするものであって、静止画コンテンツの制作時あるいは、静止画コンテンツを制作した直後更には、制作時とは無関係な時刻において上記制作された静止画コンテンツを再生することによりメタデータとすべき情報を、音声入力でコンピュータに入力することによりメタデータを制作するシステム及び方法を提供することを目的とする。

##### 【0006】

#### 【課題を解決するための手段】

上記課題を解決するために本願発明は、制作された静止画コンテンツの再生手段、上記再生手段で再生された静止画映像信号を表示する映像モニタ手段、上記映像モニタ手段でオペレータが確認した制作すべきメタデータ内容をオペレータの発声によりマイクで収録する音声入力手段、上記音声入力手段により入力された音声信号を認識する音声認識手段、音声認識手段で認識された音声情報をメタデータに変換することによりメタデータを生成するメタデータ生成手段、および上記静止画コンテンツと上記メタデータとを関連づけるため、上記静止画コンテンツに付与されている静止画番地情報と上記メタデータとを入力し番地情報付メタデータとする番地付与手段、上記制作された静止画コンテンツおよび上記番地情報付メタデータとを記録する静止画コンテンツ・メタデータ記録手段とを備えたものである。

#### 【0007】

これにより、従来キーボードで入力し、制作していたメタデータを、音声認識を用いて音声入力することにより極めて容易に制作することが可能となる。

#### 【0008】

##### 【発明の実施の形態】

本発明の請求項1に係る発明は、静止画コンテンツに関連するメタデータの制作装置であって、音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記静止画コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換し、かつ、上記静止画コンテンツとともに上記メタデータを、記録手段を用いて記録することを特徴とするメタデータ制作装置である。

#### 【0009】

本発明の請求項2に係る発明は、コンテンツに関連するメタデータの制作装置であって、音声入力手段と音声認識手段とメタデータ制作手段と上記静止画コンテンツに関連した辞書とを具備し、上記静止画コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて認識し、かつ、上記静止画コンテンツ

とともに上記メタデータを、記録手段を用いて記録することを特徴とするメタデータ制作装置である。

#### 【0010】

本発明の請求項3に係る発明は、上記請求項1ないし上記請求項2に係る発明のメタデータ制作装置であって、上記静止画コンテンツに付与されている番地情報と上記メタデータとを入力し、番地情報付メタデータを生成する番地付与手段を備え、上記静止画コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1または2のいずれかに記載のメタデータ制作装置である。

#### 【0011】

以下、本発明の実施の形態について図面を用いて説明する。

##### (実施の形態1)

図1は、本発明の実施の形態1によるメタデータ制作装置の構成を示すブロック図である。図1において、1はカメラ、2は静止画コンテンツ記録および番地付与・記録手段、3は静止画コンテンツ・メタデータ記録手段、4は静止画コンテンツ・メタデータ再生手段、5は静止画コンテンツ・メタデータ表示手段、6はマイク、7は音声認識手段、8はメタデータ生成手段、9は番地付与手段、10は辞書である。

カメラ1で撮影された静止画コンテンツは、静止画コンテンツ記録および番地付与・記録手段2に供給される。ここで、上記静止画コンテンツは記録媒体（図示せず）に記録されるとともに番地が付与され、上記番地も上記記録媒体（図示せず）に記録される。なお、上記記録媒体は、半導体メモリで構成されるのが一般的であるが、半導体メモリに限定されず、たとえば、磁気メモリ、光記録媒体、光磁気記録媒体など、種々の記録媒体を用いることが可能である。ここで記録された静止画コンテンツは出力端子201、入力端子301を介して、静止画コンテンツ・メタデータ記録手段に供給される。同様に、番地情報も出力端子202、入力端子302を介して、上記静止画コンテンツ・メタデータ記録手段3に供給される。さらに上記番地情報は上記出力端子202、入力端子902を介してメタデータ番地付与手段9（後出）に供給される。

一方、上記カメラ1で撮影された静止画に関連する情報がマイク6を介して音声

認識手段 7 に入力される。上記静止画に関連する情報は、たとえば、タイトル、撮影の日時、撮影者、撮影場所（どこ）、被撮影者（だれ）、被撮影物（なに）・・・など、撮影されたコンテンツに関連する情報である。

また、必要に応じて、音声認識用の辞書 10 も、音声認識手段 7 に供給される。音声認識手段 7 で認識された、音声データはメタデータ生成手段 8 に供給され、メタデータあるいはタグに変換される。なお、一般には、メタデータとは、このようなタグ[タイトル、撮影の日時、撮影者、撮影場所（どこ）、被撮影者（だれ）、被撮影物（なに）・・・]の集合体を意味するものである。このようにして、生成されたメタデータあるいはタグは、静止画コンテンツ自身の内容あるいはシーンとの関係を一致させるため、メタデータへの番地付与手段 7 に供給される。上記メタデータへの番地付与手段 7 にて、出力端子 202、入力端子 903 を介して供給された番地情報が上記メタデータに付与される。この様にして番地の付与されたメタデータは、出力端子 903、入力端子 303 を介して、上記静止画コンテンツ・メタデータ記録手段 3 に供給される。

静止画コンテンツ・メタデータ記録手段 3 では、同一の番地の静止画コンテンツと同一の番地のメタデータとが関連付けられて記録される。

より具体的に説明するため、上記静止画コンテンツ・メタデータ記録手段 3 にて記録された静止画コンテンツおよびメタデータを静止画コンテンツ・メタデータ再生手段 4 にて再生し、静止画コンテンツ・メタデータ表示手段 5 にて表示された結果の一例を、図 2 に示す。

図 2 において、静止画コンテンツ・メタデータ表示手段 5 の画面は、たとえば静止画コンテンツの表示部 501、番地の表示部 502 およびメタデータの表示領域 510 で構成される。メタデータの表示領域 510 は、たとえば 1) タイトルの記述部 511、2) 日時の記述部 512、3) 撮影者の記述部 513、4) 撮影場所の記述部 514、・・・などで構成される。これら、1) タイトルの記述部 511、2) 日時の記述部 512、3) 撮影者の記述部 513、4) 撮影場所の記述部 514、・・・などが、上述した音声認識により生成されたメタデータである。

上述した説明では、メタデータの生成を、静止画コンテンツの撮影の事前、撮影



とほぼ同時期あるいは、撮影の直後など、必ずしも、撮影された静止画コンテンツの確認を必要としなかった場合についての説明である。

次には、たとえば静止画コンテンツの後付として、メタデータの生成を行うなど、静止画コンテンツを再生し、モニタ手段でモニタした上記静止画コンテンツに対し、メタデータを生成する場合について図3を用いて説明する。なお、図1と同様の機能については、その説明を省略する。

カメラ1で撮影された静止画コンテンツは、静止画コンテンツ記録および番地付与・記録手段2に供給される。ここで、上記静止画コンテンツは記録媒体（図示せず）に記録されるとともに番地が付与され、上記番地も上記記録媒体（図示せず）に記録される。このような記録媒体（図示せず）を、静止画コンテンツ・番地再生手段11に供給する。静止画コンテンツ・番地再生手段11で再生された静止画コンテンツは、モニタ手段12に供給される。同様に再生された番地情報は、出力端子112、入力端子902を介して、メタデータの番地付与手段9に供給される。マイク6に向かって、音声入力する担当者（図示せず）は、上記モニタ手段12に映出された静止画コンテンツを確認のうえ、メタデータ生成に必要な言葉を発声する。このようにして、上記カメラ1で撮影された静止画に関連する情報がマイク6を介して音声認識手段7に入力される。上記静止画に関連する情報は、たとえば、タイトル、撮影の日時、撮影者、撮影場所（どこ）、被撮影者（だれ）、被撮影物（なに）・・・など、撮影されたコンテンツに関連する情報である。これ以降は、図1の説明と同様である。

#### 【0012】

なお、一般的には、音声認識には何らかの誤認識が生じる可能性がある。誤認識が生じた場合、制作されたメタデータ、タグをコンピュータ手段などの情報処理手段を用いて修正することは可能である。

#### 【0013】

##### 【発明の効果】

以上説明したように発明は、静止画コンテンツに関連したメタデータの作成あるいはタグ付けを行うに当たり、音声入力による音声認識を用い、且つ、上記メタデータあるいはタグと静止画コンテンツとの番地あるいはシーンとの関連付け

を行うため、従来のようなキーボード入力と比較して効率的にメタデータの作成やタグ付けを実施することができる。

### 【図面の簡単な説明】

#### 【図1】

本発明の実施形態1に係るメタデータ制作装置の構成を示すブロック図

#### 【図2】

本発明の静止画コンテンツ・メタデータ表示手段の一例を示す図

#### 【図3】

本発明の実施形態2に係るメタデータ制作装置の構成を示すブロック図

### 【符号の説明】

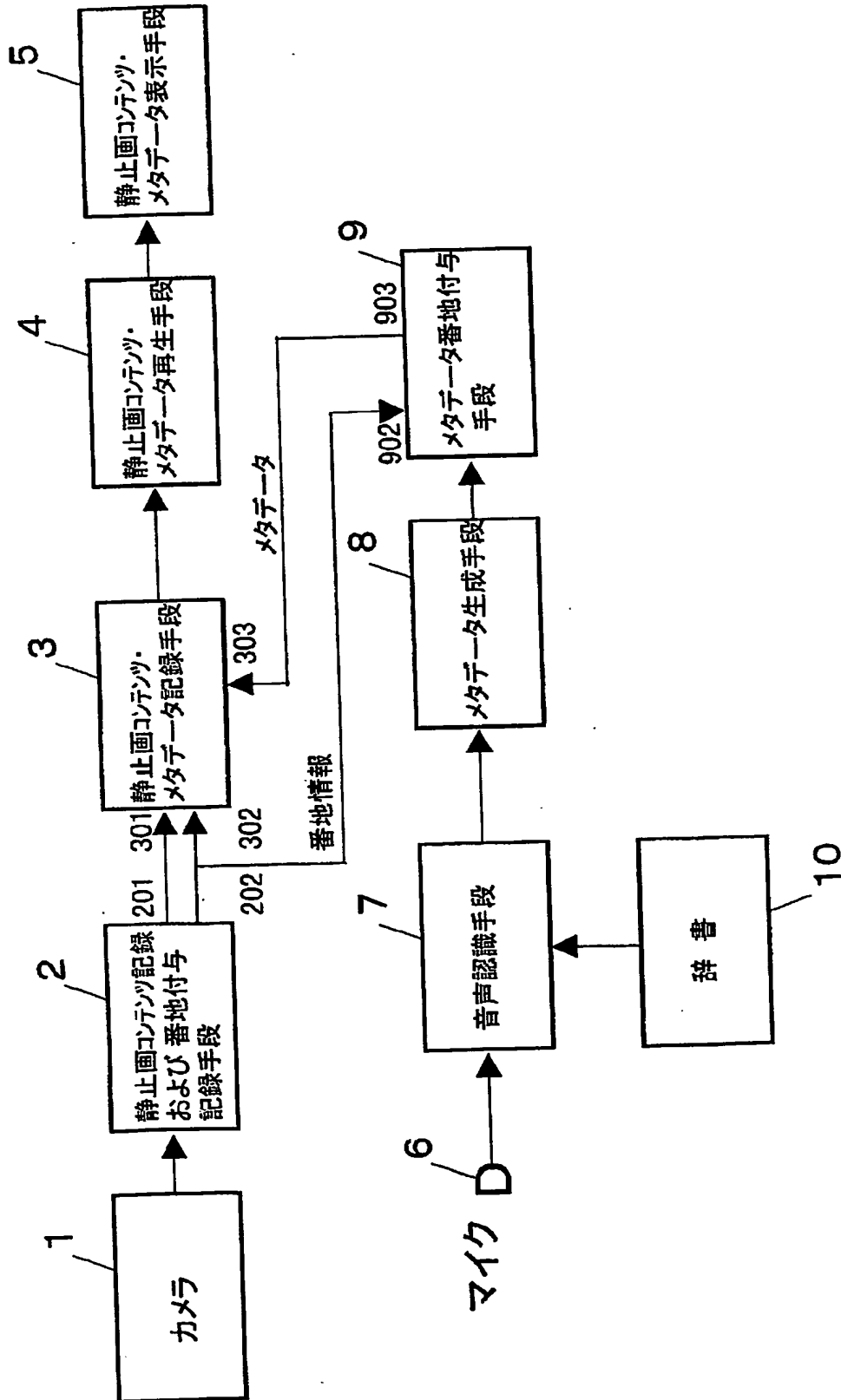
- 1 カメラ
- 2 静止画コンテンツ記録および番地付与手段
- 3 静止画コンテンツ・メタデータ記録手段
- 4 静止画コンテンツ・メタデータ再生手段
- 5 静止画コンテンツ・メタデータ表示手段
- 6 マイク
- 7 音声認識手段
- 8 メタデータ生成手段
- 9 メタデータ番地付与手段
- 10 辞書
- 11 静止画コンテンツ・番地再生手段
- 12 モニタ手段
- 111 静止画コンテンツ出力端子
- 112 番地出力端子
- 201 静止画コンテンツ出力端子
- 202 番地出力端子
- 301 映像入力端子
- 302 番地入力端子
- 303 メタデータ・番地入力端子

- 501 静止画コンテンツ表示手段
- 502 番地表示手段
- 510 メタデータ表示領域
- 511 タイトル表示領域
- 512 日時表示領域
- 513 撮影者表示領域
- 514 場所表示領域
- 902 番地入力端子
- 903 メタデータ・番地出力端子

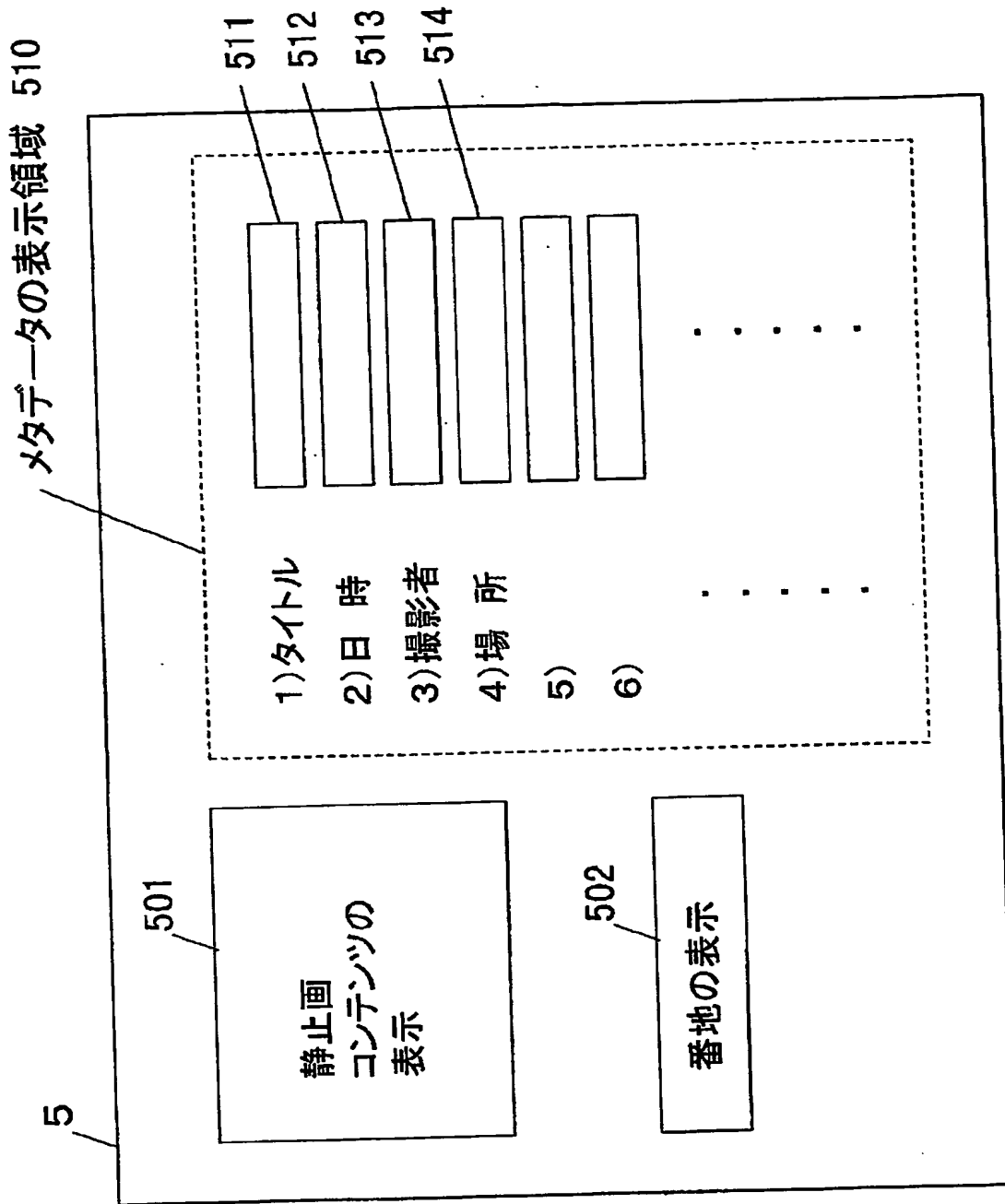
【書類名】

図面

【図 1】

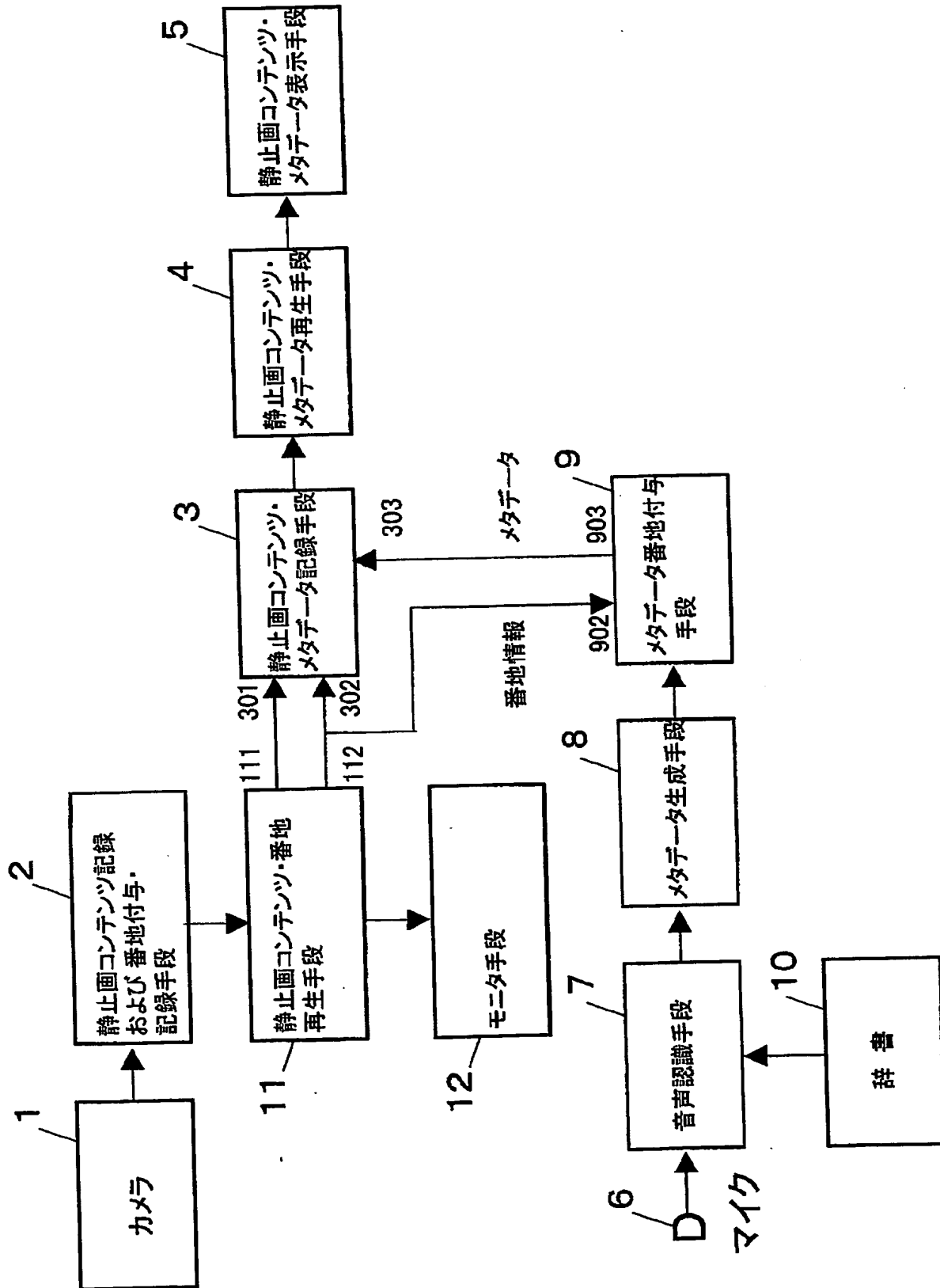


【図 2】



静止画コンテンツ・メタデータ表示手段での表示例

【図 3】



【書類名】 要約書

【要約】

【課題】 静止画コンテンツ制作において、これらコンテンツに関連したメタデータの付与するに際し、付与する際の労力を低減することを目的とする。

【解決手段】 音声入力手段と、音声認識手段と、メタデータ生成手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ生成手段によりメタデータに変換する。静止画コンテンツに付与されている静止画番地情報と上記メタデータとを入力し番地情報付メタデータとする番地付与手段、上記制作された静止画コンテンツおよび上記番地情報付メタデータとを記録する静止画コンテンツ・メタデータ記録手段とを備える。

【選択図】 図1

特願 2002-319757

出願人履歴情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社